DOCUMENT RESUME

ED 438 297 TM 030 605

AUTHOR Morris, Michael; Jacobs, Lynette
TITLE You Got a Problem with That? Exploring Evaluators' Disagreements about Ethics.
PUB DATE 1999-11-05
NOTE 39p.; Paper presented at the Annual Meeting of the American Evaluation Association (Orlando, FL, November 2-6, 1999).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Ethics; *Evaluation Methods; Evaluation Utilization; *Evaluators; Moral Values

ABSTRACT

Research has suggested that evaluators vary in the extent to which they interpret the challenges they face in ethical terms. The question of what accounts for these differences was explored through a survey completed by 391 individuals listed in the database of the American Evaluation Association. The first section of the questionnaire presented scenarios of evaluators in situations with potential ethical issues and asked respondents to rate the evaluator's actions. The second section asked respondents to rate the usefulness of four role-oriented labels--consultant, scientist, reporter, and facilitator--for describing the work that evaluators do. The third section asked for demographic data about the respondents. The most striking finding related to the scenarios was the lack of consensus that characterized respondents' judgments of whether each hypothetical evaluator had behaved ethically. It was possible to identify subgroups of responses, and it was apparent that respondents in private business and consulting were most likely to criticize the evaluator's behavior on ethical grounds. Respondents' judgments about ethics were related only to their view of the consultant role; the more useful that role was perceived to be, the less likely the respondent was to view the evaluator's actions as ethically problematic. Although there was little agreement in respondents' views, the more information the respondents had, the more likely they were to agree on what the evaluator was ethically obliged to do. An appendix describes the scenarios. (Contains 10 tables and 17 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 438 297

You Got a Problem with That?

Exploring Evaluators' Disagreements about Ethics

Michael Morris and Lynette Jacobs

University of New Haven

M. Morris

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paper presented at the annual meeting of the American Evaluation Association, Orlando, FL,

November 5, 1999

Address correspondence to: Michael Morris, Ph.D.
Department of Psychology
University of New Haven
West Haven, CT 06516

203-932-7289

mikepsyc@charger.newhaven.edu.

TM030605

BEST COPY AVAILABLE

You Got a Problem with That?

Exploring Evaluators' Disagreements about Ethics

Ethical issues in evaluation have received increasing attention in recent years (e.g., Fitzpatrick and Morris 1999; House and Howe 1999; Morris 1999b; Newman and Brown 1996; Shadish, Newman, Scheirer, and Wye 1995; Wenger, Korenman, Berk, and Liu, 1999). Not surprisingly, one outcome of this attention has been a recognition of the diverse perspectives which evaluators bring to the domain of ethics. Indeed, considerable disagreement even appears to surround such basic questions as, *What constitutes an ethical issue in evaluation*? For example, when summarizing their research on evaluation ethics, Newman and Brown (1996, 89) note tnat "we consistently found people whose generalized response was 'What? Ethics? What does ethics have to do with evaluation?' This came from experienced evaluators, long-term users of evaluation, evaluation interns, and faculty members teaching program evaluation." In a similar vein, Morris and Cohn (1993) found that 35% of their sample of American Evaluation Association (AEA) members responded "no" when asked in a questionnaire, "In your work as a program evaluator, have you ever encountered an ethical problem or conflict to which you had to respond?" Finally, in an interview study whose goal was to identify and describe the ethical issues encountered by public-sector evaluators, Honea (1992, 317) found that "ethics was not discussed during the practice of evaluation and ethical dilemmas were rarely, if ever, identified during the conduct of evaluation and policy analysis activities."

Taken as a whole, these results suggest that evaluators vary in the extent to which they "interpret the challenges they face in ethical terms" (Morris, 1999a, 16). What accounts for these differences? The present study explores this question by analyzing evaluators' responses to a set

of detailed scenarios drawn from professional practice. Scenarios, because they are specific and concrete, are more likely than open-ended methods (e.g., Honea, 1992; Morris and Cohn, 1993) to generate uniform reference points for the application of respondents' opinions, beliefs, and values related to ethics. This, in turn, increases the likelihood that observed differences in respondents' views represent real, substantive differences which have practical implications, a conclusion that is harder to justify when disagreements pertain to issues presented in more abstract, theoretical terms.

Previous research has identified several factors that might influence an individual's tendency to perceive evaluation problems through an "ethical lens." Morris and Cohn (1993) found that evaluators who reported that they had *never* encountered an ethical conflict in their work had conducted fewer evaluations, had devoted more of their time to internal evaluation, and were more likely to have been trained in the field of education than respondents who said that they *had* encountered such challenges. Also relevant in this context are two factors identified by Honea (1992) on the basis of her interviews with public-sector evaluators: allegiance to the role of *objective scientist*, and membership in an evaluation *team*. Honea believes that internalization of the scientist role, and participation in research teams, decreases the extent to which one sees ethical issues -- as opposed to methodological or political ones -- as salient in one's evaluation work. In the current investigation, we attempt to examine with greater directness and precision the role played by these and other factors in the perceptions of challenges that might be deemed ethical in nature.

Method

Participants

The population for the study consisted of the 3167 individuals with US addresses who were listed in the March 1999 database of AEA. We mailed a questionnaire to a random sample of 798 of these individuals. A small number of surveys (24) were returned due to incorrect addresses, reducing the original sample size to 774. Overall, we received 397 responses, which represents a return rate of 51%. Within this group there were 6 individuals who indicated that they were not evaluators or evaluators-in-training, and thus they did not think it was appropriate for them to complete the survey. Consequently, the data analyses reported here are based on a sample of 391.

Survey Instrument

The questionnaire contained three sections. Section A included three, single-paragraph scenarios (see Appendix); in each one an evaluator acts in a way that could be deemed ethically problematic. In the first scenario, hereafter referred to as the "Revised Report," the evaluator alters a section of a final report in response to pressure from a stakeholder. In the second scenario (Advisory Group), an evaluator assembles a widely representative advisory group for a project, but does not *actively* involve these stakeholders in the evaluation process. In the third scenario (Passive Consent), the evaluator decides to use passive rather than active consent when studying a school-based youth program, even though he/she realizes that "some parents who oppose the research will simply forget to return the passive-consent form, while others who *would have* been opposed to the study will fail to read it in the first place." Because three scenarios can be sequenced in six different ways, there were six versions of the questionnaire, with each version

(representing a different sequence) accounting for one-sixth of the total number of surveys mailed.

On a Likert scale, respondents indicated the extent to which they regarded the evaluator's actions in each scenario as "ethically problematic" (1 = they definitely *are* problematic, 2 = they probably *are* problematic, 3 = unsure, 4 = they probably are *not* problematic, 5 = they definitely are *not* problematic). A note at the beginning of the survey encouraged respondents to define "ethical" in terms of "issues of morality, i.e., good and bad, right and wrong, duty and obligation." For each scenario, respondents explained, in an open-ended fashion, *why* they gave the answer they did to the Likert item. Finally, we asked respondents to predict (by assigning percentages to the five Likert categories) how the overall AEA membership would react to each scenario.

In the survey's second section, respondents rated on a seven-point scale the usefulness of four role-oriented labels -- consultant, scientist, reporter, and facilitator -- for describing the work that evaluators do (1 = not at all useful, 7 = extremely useful). Respondents also rated on a five-point scale the usefulness of AEA's Guiding Principles for Evaluators (1= not at all useful, 5 = extremely useful), and indicated their overall political orientation (1 = very conservative, 7 = very liberal).

The final section of the questionnaire solicited background information. Respondents reported the number of years they had worked in evaluation, as well as the approximate number of evaluations they had conducted. They also estimated the percentage of evaluations they had conducted in each of the following capacities: external evaluator, internal evaluator, member of an evaluation team, and solo practitioner. In addition, information was gathered on the respondent's highest degree, primary discipline, employment setting, and sex.

Results

Respondent Characteristics

A majority of the respondents possessed a doctoral degree (54% Ph.D., 7% Ed.D.); 30% had a master's, 5% a bachelor's, and 5% were "other." The primary discipline of over half the respondents was either education (19%), psychology (17%), or evaluation (17%) (see Table 1).

The largest subgroup of respondents worked in a college or university (40%), with private business/consulting (19%) and non-profit organizations (15%) representing the only other settings employing 10% or more of the sample (see Table 2). With respect to sex, 52% of the respondents were female, and 48% were male. This sex ratio differed significantly from that of the nonrespondent group, where 60% were female and 40% were male (χ^2 [1, N = 79] = 4.52, $p < .05$). However, we found the respondent's sex to be unrelated to the key variable examined in this study (i.e., reactions to the three scenarios), and thus we have little reason to believe that the different response rates for males and females affected our results in a substantive way.

Evaluation Experience

Respondents had worked in the evaluation field for an average of 12.5 years (SD = 9.2), with a range spanning from 0 to 52 years. Over half (53%) had conducted 11 or more evaluations (see Table 3). Both external and internal evaluators were well represented in the sample, as were team evaluators and solo practitioners (see Table 4). At the extremes, purely external evaluators accounted for 28% of the respondents, while purely internal ones accounted for 13%. Those who had only participated in team evaluations comprised 20% of the respondents, while those who always worked alone represented 7%.

Reactions to the Scenarios

Respondents' views of the scenarios are presented in Table 5. The evaluator's actions were seen as most troubling in the Passive Consent vignette, with 69% of the sample rating the use of passive consent in this situation as definitely or probably ethically problematic. Just over 50% of the respondents regarded the behavior described in the Revised Report scenario as ethically problematic, while only 39% believed that the evaluator's failure to involve stakeholders actively in the Advisory Group scenario was problematic.

Respondents' predictions of how the AEA membership would react to the scenarios were strongly related to their own views of the vignettes.[1] The more convinced a respondent was that the evaluator had behaved unethically in a given scenario, the larger the respondent's estimate of how many AEA members would view the evaluator's action as ethically problematic (see Table 6). For example, when respondents viewed the evaluator's actions as definitely or probably ethically problematic, their mean estimate of the percentage of the AEA membership that would share this view was 69%. In contrast, when respondents regarded the evaluator's behavior as *not* problematic (definitely or probably), they estimated that only 28% of AEA would consider the evaluator's actions to be ethically problematic (definitely or probably).

Content Analysis

For the purpose of content-analyzing the respondents' explanations of their answers, we grouped respondents into three categories for each scenario: those who thought the evaluator's actions were definitely or probably ethically problematic; those who were unsure whether the evaluator's actions were ethically problematic; and those who thought the evaluator's actions were definitely or probably *not* ethically problematic. In this section we focus on the explanations

offered by the first and third groups; the "Unsures" are omitted.

As might be expected, the specific issues raised by respondents differed across the three scenarios. Given the study's focus on ethics, we used the Guiding Principles for Evaluators (American Evaluation Association, 1995) -- Systematic Inquiry, Competence, Integrity/Honesty, Respect for People, and Responsibilities for General and Public Welfare -- as a conceptual tool for categorizing these open-ended responses, once we had conducted an initial content analysis to identify specific themes in the explanations. The results indicate that a general principle (e.g., Integrity/Honesty) could support arguments both for and against the ethicality of the evaluator's actions in a given scenario (see Tables 7, 8, and 9). For example, 30% of those who faulted the evaluator in the Advisory Group scenario maintained that extensive stakeholder participation is required for an accurate evaluation. In contrast, 11% of those who found the evaluator's actions in that scenario to be acceptable believed that such participation could jeopardize the evaluation's objectivity. Both of these arguments pertain most directly to the principle of Systematic Inquiry ("Evaluators should adhere to the highest appropriate technical standards in conducting their work.....so as to increase the accuracy and credibility of the evaluation information they produce" [American Evaluation Association, 1995, 22]).

In other cases, a principle's ability to encompass conflicting arguments was related to respondents' interpretations of a lack of detail in the scenario. Thus, 61% of those who objected to the evaluator's behavior in the Revised Report scenario assumed that the revision substantially altered the report, while 70% of those who did *not* object gave explanations which indicated that they did not share this assumption (see Table 8). In both instances, the relevant principle involves the Integrity/Honesty of the evaluation ("Evaluators should not misrepresent their procedures,

data, or findings" [American Evaluation Association, 1995, 23]).

Finally, each scenario generated a certain number of open-ended responses which claimed that the situation depicted did not raise an ethical issue. Indeed, among those who saw the evaluator's actions as not problematic in the Advisory Group scenario, 50% thought the case involved a methodological or philosophical issue, **not** an ethical one. The percentages of the not-problematic subgroup who believed this in the other two scenarios were much smaller (4%-5%).

Combined Scenarios

The analyses in this section group respondents into three categories: those who responded "definitely problematic" or "probably problematic" to *none* of the three scenarios (8% of the sample); those who responded in this fashion to *one or two* of the scenarios (76% of the sample); and those who found the evaluator to be at fault (definitely or probably) in *all three* of the scenarios (16% of the sample). We used either one-way ANOVA or Chi-square tests to examine the relationship of this variable to responses to questions dealing with evaluator role, the Guiding Principles for Evaluators, political orientation, evaluator experience, educational/employment background, and sex (see Table 10).

The only role for which a significant relationship was found was consultant: viewing all three scenarios as problematic was *negatively* associated with believing that the consultant label is useful for describing the work of evaluators. In contrast, viewing the scenarios as problematic was *positively* associated with finding the Guiding Principles for Evaluators to be useful. Other significant relationships included the following:

-- Respondents employed in private business/consulting were less likely than those in other settings to believe that the scenarios involved ethically problematic behavior on the

evaluator's part.

-- Among those *not* employed in private business/consulting, length of evaluation experience (as measured in terms of both years and number of evaluations conducted), was *negatively* related to judging the evaluator's actions to be ethically problematic. No such relationship characterized those who *were* employed in private business/consulting.

-- Among those employed in private business/consulting, respondents who viewed all of the scenarios as problematic devoted less of their time to external evaluation than those who judged none, or one or two, of the scenarios as problematic. This relationship was not found among those employed in other settings.

Holders of a doctoral degree were less likely than BA/MA respondents to see all of the scenarios as ethically problematic. However, this relationship is attributable to the greater evaluation experience of the former group, in terms of both years and number of evaluations conducted. When either of these experience indicators is held constant, the relationship between degree and one's score on the combined scenarios disappears.

The Not-an-Ethical-Issue Subgroup

Only the Advisory Group scenario produced enough open-ended explanations (106) of "I don't think [or "I'm not sure"] this is an ethical issue" to warrant further statistical analysis. When we compared this subgroup with the rest of the sample on the variables examined in the previous section (evaluator roles, Guiding Principles for Evaluators, etc.), no significant differences emerged.

Discussion

The results of this study shed light on two important, and related, questions in evaluation

ethics. First, what issues do evaluators emphasize, and disagree about, when judging the ethicality of professionals' behavior in specific situations? Second, are there factors which operate at a *more general level* to increase or decrease the salience of ethical concerns in the eyes of evaluators? We will address both of these questions in this section.

Consent, Reporting, and Stakeholder Participation

Perhaps the most striking finding pertaining to the individual scenarios is the lack of consensus which characterized the respondents' judgments of whether or not each of the hypothetical evaluators had behaved ethically. Even in the Passive Consent scenario, where agreement was highest, only 44% of the respondents believed that the evaluator's actions were *definitely* problematic.

In part, widespread disagreement may simply reflect the limitations of scenario methodology. A single-paragraph vignette inevitably leaves many details unspecified, and different respondents are likely to "fill in the blanks" with different assumptions, with some of these assumptions having implications for the ethical judgments rendered. Thus, as was previously mentioned, respondents to the Revised Report scenario varied in their views of how the evaluator's revisions would influence the report: of those who found the evaluator's behavior ethically problematic, 61% cited the inappropriateness of substantively altering a fair report; of those who were unsure of the behavior's ethicality, 66% said they were unsure *because* they did not know if the revisions substantively altered a fair report; and among those who saw the evaluator's actions as not problematic, 70% assumed the revisions had *not* misrepresented the study's key findings.

Similar points could be made concerning the other two vignettes. In the Advisory Group

scenario, it appears that respondents varied in their views of the understanding established between the evaluator and the stakeholders at the beginning of the project concerning the nature of the "advisory" relationship. And in the Passive Consent scenario, respondents differed in the extent to which they assumed that the "school-based youth program" involved high-risk issues. Had the scenarios been more explicit about these and other issues, it is likely that respondents would have displayed higher levels of agreement when judging the evaluator's behavior in each vignette.

Reducing disagreement is not synonymous with *eliminating* it, however. Even if the Revised Report scenario had contained actual copies of the both the original and final reports, respondents would almost certainly have differed over whether the changes in the document represented "substantive" ones or not, resulting in conflicting conclusions about the ethicality of the evaluator's actions. The same principle applies to the other two scenarios: describing more fully the initial evaluator-stakeholder conversations in the Advisory Group vignette, and specifying the type of youth program in the Passive Consent scenario, does not guarantee that respondents would have agreed on the *nature* of the understanding in the former vignette, or the *amount* of risk involved in the latter one. Indeed, as Korenman, Berk, Wenger, and Lew (1998, 47) observe, "ambiguity is.....typical of real-life behaviors as well as scenarios." With these considerations in mind, we are inclined to conclude that the level of disagreement among our respondents on the issues raised in the three scenarios is probably *less* than the reported percentages suggest, but of considerable magnitude nonetheless.

Both the Passive Consent and Revised Report scenarios were apparently seen by nearly all respondents as encompassing ethical problems. Only 1% of the sample, when explaining their

judgments of the evaluators' actions, claimed that these vignettes did not raise an ethical issue. In contrast, 19% of the respondents expressed such an opinion when discussing the Advisory Group scenario. Why the difference? The Passive Consent scenario deals with informed consent, traditionally a core topic in discussions of research ethics (e.g., Newman and Brown 1996, 147-149). Similarly, at the heart of the Revised Report scenario is the issue of impartial reporting of findings, a professional responsibility that researchers typically see as having major ethical significance (Korenman et al. 1998; Morris and Cohn 1993). The Advisory Group scenario, however, focuses on stakeholder involvement and empowerment, a domain that in the minds of many evaluators does not necessarily suggest a set of *ethical* imperatives. "Empowerment Evaluation" (Fetterman, Kaftarian, and Wandersman 1996), for example, has been the source of considerable controversy within the field (see Fetterman 1997; Patton 1997; Scriven 1997). Thus, it should not be surprising that, of the respondents who did not see the evaluator's actions as unethical in this scenario, 50% indicated that they did not believe the problem involved was an ethical one. Representative comments from this subgroup included, "Involving stakeholders is a matter of use not ethics," "This may not be the *smartest* approach, but I don't find it an ethical dilemma," and "While not actively involving stakeholders is not good evaluation, I don't see it as 'morally' wrong." As previously reported, these respondents did not significantly differ from the rest of the sample on any of the variables examined in the study.

When the explanations respondents offered for their ethical judgments of the three scenarios are viewed as a whole, the differences between them reflect a dynamic commonly found in controversy: conflicting views of whether a *general* principle or value is being upheld in a *specific* situation. Thus, respondents who found fault with the evaluator's behavior in the Passive

Consent scenario usually thought that the spirit of informed consent (if not the "letter of the law") had been violated by the evaluator. In contrast, most of those who were ethically comfortable with this scenario indicated that they did not see the evaluator's actions threatening informed consent. Virtually all of the respondents would probably claim that their positions were consistent with the AEA Guiding Principle of Respect for People. Likewise, in most cases both the defenders *and* critics of the evaluator in the Revised Report scenario professed their allegiance to the importance of not altering the substance of the final report, and undoubtedly saw themselves upholding the Integrity/Honesty of the evaluation. And in the Advisory Group vignette, there was one subgroup of respondents who argued that an accurate evaluation *required* extensive stakeholder participation, while another subgroup claimed that such participation would *threaten* the evaluation's accuracy. Both groups would probably maintain that they were committed to Systematic Inquiry in the evaluation.

These findings underscore one of the limitations of any set of highly general set of principles for guiding professional behavior (e.g., House 1995; Rossi 1995). As Rossi (1995, 59) has observed of such principles, "I am certain that I can claim to subscribe to them. I am also certain that if I held very different views of evaluation, I would also be in compliance."

Who Finds Fault, and Who Doesn't?

Although in this study we failed to identify a distinctive respondent subgroup whose *general* orientation was an explicit one of not viewing evaluation problems through an "ethical lens," we did succeed in generating a composite variable (Combined Scenarios) that may reflect a similar orientation operating at a more implicit level. Specifically, we distinguished between three groups of respondents: those who believed that the evaluator's behavior was ethically problematic

in *none* of the scenarios; those who found it problematic in *one or two* of the scenarios; and those who faulted the evaluator in *all three* of the scenarios. In at least one crucial respect, perceiving an evaluator's actions as ethically blameless is much the same as perceiving the evaluator's behavior as not involving an ethical issue: in neither case is a judgment of moral wrongdoing rendered.

When respondents were subgrouped in this fashion, the differences that emerged between them were intriguing. Perhaps most telling was the role of primary employment setting. Respondents in private business/consulting were less likely than those in other settings to criticize ethically the evaluator's behavior. This finding underscores the importance of structural/background variables in understanding evaluators' ethical perceptions. Evaluators in private business/consulting essentially "work for themselves," a status that can be counted on to heighten one's sensitivity to the personal economic consequences of one's actions. Viewing an evaluator's behavior as ethically inappropriate usually implies that some other action should have been taken, an action that in many cases might not, at least in the short term, be in the evaluator's material self-interest. Thus, experiences in private business/consulting may predispose evaluators to be more tolerant, and "understanding," of behavior that those in other settings might criticize ethically. The influence of role-oriented variables on ethical judgments relevant to evaluation has also been documented by Korenman et al. (1998), who found that National Science Foundation research grantees were more likely than administrators responsible for academic research integrity to perceive violations of collegiality and sharing of research products as unethical.

Viewing AEA's Guiding Principles for Evaluators as "useful for thinking about the ethical issues you encounter in evaluation" was positively related to believing that the evaluators'

scenario behavior was unethical. It is unclear whether the perceived value of the Guiding Principles actually plays a *causal* role with respect to the ethical judgments participants rendered in the study. It may be that ethical issues have greater salience for some individuals than others, and this salience causes the former group to find the Guiding Principles more useful, *as well as* to be more critical of evaluators' behavior with respect to ethics.

Interestingly, simply having *knowledge* of the Guiding Principles does not appear to be important in this regard: those who responded that they were not familiar with the Guiding Principles (48% of the sample) were no more likely than those who were familiar with them to perceive unethical behavior in the scenarios. This finding lends support to the notion that a causal factor other than the Guiding Principles is responsible for the observed relationship between the Guiding Principles' subjective value and reactions to the scenarios.

Of the four roles examined in this study -- consultant, scientist, reporter, and facilitator -- respondents' ethical judgments were only related to their view of the consultant role: the more useful that role was perceived to be, the less likely the respondent was to view the evaluators' actions as ethically problematic. The nature of the consultation process may be key to interpreting this result. A consultant is typically defined as an expert "who gives professional advice or services" (Webster's Ninth New Collegiate Dictionary 1988, 282). Inherent in this view is the notion that, within their domain of expertise, the judgments reached by consultants are worthy of respect or trust. Indeed, it is precisely for these judgments that consultants are hired by clients in the first place. Thus, respondents who highly value the consultant role for evaluation may be signaling, in part, a willingness to give evaluators the "benefit of the doubt" when scrutinizing their behavior in specific situations. Such an orientation could lead to fewer

accusations of unethical behavior than would otherwise be the case.

In this context, the failure to find a relationship between the perceived usefulness of the scientist role and views of unethical behavior deserves comment. This role was included in the survey to test, in an admittedly limited way, Honea's (1992) conclusion that internalization of the objective scientist role decreases the salience of ethical issues for evaluators. Our finding does not support Honea's conclusion. It is possible, of course, that our operationalization of the scientist role was too limited to do justice to her conceptualization of it. It should also be noted that our study focused on respondents' perceptions of *other* evaluators' experiences (as represented by the scenarios), rather than their own. This difference in method between our study and hers may, at least in part, be responsible for the different results obtained. This factor may also help explain the lack of a relationship we found between involvement in team-oriented evaluations -- a dimension deemed important by Honea -- and perceptions of unethical behavior in the scenarios.

Among those currently employed in private business/consulting, respondents with more experience in external evaluation were less likely to see the scenarios as ethically problematic. This result may reflect the fact that -- by definition -- all evaluations conducted by those in private business/consulting are external in nature. Thus, within the private business/consulting subgroup, the percentage of external evaluations conducted probably serves as a rough proxy for *how long* a respondent has been in private business/consulting. Hence, this finding might be viewed as further evidence of the relationship between the private business/consulting role and ethical judgments.

Finally, we found that among those *not* employed in private business/consulting, evaluation experience was negatively associated with believing the evaluators' actions in the

scenarios were ethically problematic. In this regard, it should be noted that training programs in most disciplines include curricular components that address ethical issues in an explicit fashion. With this exposure relatively fresh in their minds, less experienced evaluators may be prone to "set the bar higher" for ethical decision-making than more seasoned practitioners, who have had more encounters than the former group with the myriad factors that can constrain these decisions. As one respondent with 15 years of experience wrote when defending the evaluator in the Passive Consent scenario, "[the evaluator's actions are] not ethically problematic, just realistic. The benefit of the evaluation results justifies trying to get as good a sample as possible."

At first glance, this finding for experience might be viewed as contradicting the results of the Morris and Cohn (1993) study, where experienced evaluators were more likely than the less experienced to report that they had encountered ethical conflicts in their work. Once again, however, the different focuses of the two studies are key. The Morris and Cohn investigation targeted the *respondents'* experiences, while the current study examined the respondents' reactions to *others'* experiences. As one's evaluation experience grows over time, the number of "opportunities" one has to encounter an ethical problem grows as well, which is what Morris and Cohn found. In addition, Morris and Cohn did not ask respondents to pass ethical judgment on their own behavior, while in the research reported here we did request that such judgments be rendered concerning the actions of the hypothetical evaluators.

Conclusion

To those who would like to see evaluators "speak with one voice" on ethical matters, this study delivers two messages. The bad news is that one voice does not exist, at least on the scenarios we examined involving stakeholder involvement, reporting of results, and informed

consent. The good news, based on respondents' explanations of their views, is that there is reason to believe that the more information evaluators have about a specific challenging situation, the more likely they are to agree on what the evaluator is ethically obligated to do. Although it may be true that "the devil is in the details," it is also the case that the most meaningful common ground is likely to be found there, rather than in more abstract discussions. As House (1995, 27) has put it, "ethical problems are manifested only in particular concrete cases, and endorsement of general principles sometimes seems platitudinous or irrelevant."

Of course, even with a surfeit of details, significant disagreement is likely to remain in many instances. Applying general ethical principles and standards to a particular circumstance can leave a great deal of room for value-based interpretation and differences in prioritization, as is evident from arguments over whether scientific objectivity is enhanced or hindered by extensive stakeholder involvement, to cite just one example.

Against this background, increased dialogue among evaluators who bring different orientations to ethical problems is likely to be valuable to the field, in terms of both theory and practice. Our results suggest that evaluators in private business/consulting and those in other settings would especially benefit from talking more with each other, as would new and experienced evaluators. To the extent that evaluators assume that *other* evaluators share their opinions about ethical issues -- a pattern that was striking in this study -- a fuller appreciation of the "ethical lenses" that can be applied to evaluation challenges awaits those who participate in such a dialogue.

REFERENCES

American Evaluation Association, Task Force on Guiding Principles for Evaluators. 1995. Guiding Principles for Evaluators. In *Guiding Principles for Evaluators*. Edited by W. R. Shadish, D. L. Newman, M. A. Scheirer, and C. Wye, 19-26. New Directions for Program Evaluation, no. 66. San Francisco: Jossey-Bass.

Fetterman, D. 1997. Empowerment evaluation: A response to Patton and Scriven. *Evaluation Practice* 18: 253-266.

Fetterman, D. M., Kaftarian, S. J., and A. Wandersman, eds. 1996. *Empowerment evaluation: Knowledge and tools for self-assessment and accountability*. Thousand Oaks, CA: Sage.

Fitzpatrick, J. L., and M. Morris, eds. 1999. *Current and emerging ethical challenges in evaluation*. New Directions for Evaluation, no. 82. San Francisco: Jossey-Bass.

Honea, G. E. 1992. *Ethics and public sector evaluators: Nine case studies*. Unpublished doctoral dissertation, University of Virginia.

House, E. R. 1995. Principled evaluation: A critique of the AEA Guiding Principles. In *Guiding Principles for Evaluators*. Edited by W. R. Shadish, D. L. Newman, M. A. Scheirer, and C. Wye, 27-34. New Directions for Program Evaluation, no. 66. San Francisco: Jossey-Bass.

House, E. R., and K. R. Howe. 1999. *Values in evaluation and social research*. Thousand Oaks, CA: Sage.

Korenman, S. G., Berk, R., Wenger, N. S., and V. Lew. 1998. Evaluation of the research norms of scientists and administrators responsible for academic research integrity. JAMA 279: 41-47.

Morris, M. 1999a. Research on evaluation ethics: What have we learned and why is it important? In *Current and emerging ethical issues in evaluation*. Edited by J. L. Fitzpatrick and M. Morris, 15-24. New Directions for Evaluation, no. 82. San Francisco: Jossey-Bass.

Morris, M. 1999b. Ethical Challenges. *American Journal of Evaluation* 20: 113-122.

Morris, M., and R. Cohn. 1993. Program evaluators and ethical challenges: A national survey. *Evaluation Review*, 17: 621-642.

Newman, D. L., and R. D. Brown. 1996. *Applied ethics for program evaluation*. Thousand Oaks, CA: Sage.

Patton, M. Q. 1997. Toward distinguishing empowerment evaluation and placing it in a larger context. *Evaluation Practice* 18: 147-163.

Scriven, M. 1997. Empowerment evaluation examined. *Evaluation Practice* 18: 165-175.

Shadish, W. R., Newman, D. L., Scheirer, M. A., and C. Wye, eds. 1995. *Guiding Principles for Evaluators*. New Directions for Program Evaluation, no. 66. San Francisco: Jossey-Bass.

Webster's ninth new collegiate dictionary. 1988. Springfield, MA: Merriam-Webster.

Wenger, N. S., Korenman, S. G., Berk, R., and H. Liu. 1999. Reporting unethical research behavior. *Evaluation Review*, 23: 553-570.

Appendix

The Revised Report Scenario

An evaluator has recently shared the draft of a final report with the director of the program being evaluated. After reviewing the draft, the program director asks the evaluator to "tone down" one section of the report which describes some operational problems within the program. The director believes that the findings in this section -- although accurate -- are presented in a way which could cause readers to overlook the *overall* success of the program's implementation. (The evaluator's sponsor and primary client is a philanthropic foundation which is the major source of funding for the program.) The evaluator reexamines the draft and concludes that the findings on operational problems have been reported in a fair and balanced fashion. Nevertheless, the evaluator wishes to be responsive to the director's concerns. The evaluator revises the section in question, mainly by deleting a number of harshly worded quotes concerning operational difficulties that were voiced by interview and survey respondents.

The Advisory Group Scenario

An evaluator is conducting an impact study of an urban crime prevention program. Key stakeholders include: the funding source (a local foundation); the community agency responsible for overseeing implementation of the program; the police department; the mayor's office; local merchants; neighborhood block watch groups; and several organizations specializing in youth services. The evaluator assembles an advisory group for the evaluation which includes representatives from all of these constituencies. As the project unfolds, the evaluator mainly uses the advisory group meetings to keep stakeholders informed of the evaluation's progress. The evaluator places very little emphasis on actively involving stakeholders in the process of

conceptualizing the evaluation and how it should be carried out, or in interpreting the data. The evaluator's experience in doing research on crime prevention interventions significantly exceeds that of any of the stakeholders.

The Passive Consent Scenario

In evaluating a school-based youth program, the evaluator has the choice of using either an active-consent or a passive-consent procedure to obtain parental permission. Active consent requires parents to sign and return a form if they wish to give permission for their child to participate in a study. In contrast, passive consent only requires them to sign and return a form if they do *not* want their child to participate. In general, it is much easier to achieve high participation rates with passive-consent approaches than active-consent ones. In this particular situation, the evaluator is convinced that passive consent will generate a significantly higher participation rate than active consent, and be much less costly to implement as well. The evaluator believes that, in part, this higher rate will result from the fact that some parents who oppose the study will simply forget to return the passive-consent form, while others who *would have* opposed the study will fail to read the form in the first place. The evaluator decides to use the passive consent procedure.

TABLE 1: Primary Discipline of Respondents (n = 390)

Discipline	*Percentage*
Education	19
Psychology	17
Evaluation	17
Public Admin./Pol. Science	10
Research/Statistics	9
Sociology	7
Social Work	4
Public Health	3
Other	14

TABLE 2: Employment Settings of Respondents (n = 391)

Setting	*Percentage*
College/University	40
Private Business/Consulting	19
Nonprofit organization	15
Federal agency	9
State agency	5
Local agency	3
School system	3
Other	6

TABLE 3: Evaluation Experience of Respondents (n = 390)

Number of Evaluations Conducted	*Percentage*
None	4
1-5	24
6-10	19
11-19	16
20 or more	37

TABLE 4: Experience in External and Team Evaluations (n = 373-374)

Percentage of Evaluations Conducted in a Given Role	*Evaluator Role*	
	External Evaluator	*Team Evaluator*
76-100	46	43
50-75	17	24
25-49	8	12
0-24	29	20

Note. Values in the two right-hand columns represent the percentage of respondents in a given role category. Figures for Team Evaluator do not total 100% due to rounding.

TABLE 5: Reactions to the Three Scenarios (in percentages) (n = 391)

	Scenario		
Were the Evaluator's Actions Ethically Problematic?	*Passive Consent*	*Revised Report*	*Advisory Group*
Definitely *are* problematic	44	23	19
Probably *are* problematic	25	28	20
Unsure	11	17	12
Probably are *not* problematic	16	28	32
Definitely are *not* problematic	4	4	17

TABLE 6: Predictions of AEA Members' Scenario Judgments as a Function of One's Own Judgments (n = 297-304)

	Predicted AEA Judgment		
Respondent's Judgment	*Problematic*	*Unsure*	*Not Problematic*
Evaluator's behavior is definitely or probably ethically problematic	69	12	19
Unsure	35	35	30
Evaluator's behavior is definitely or probably *not* ethically problematic	28	16	56

Note. Values represent respondents' mean predicted percentages of the AEA membership who would judge a scenario in a given way (all three scenarios combined).

TABLE 7: Respondents' Explanations: The Passive Consent Scenario

Respondents Who Judged the Evaluator's Actions as Definitely or Probably Ethically Problematic (n = 271)

Percentage	Explanation	Relevant Guiding Principle
45	The evaluator is *consciously* violating informed consent by using passive consent despite his/her knowledge of its limitations in this situation	Respect for People
22	Passive consent is not permitted under various legal/policy guidelines	Respect for People
15	Passive consent is inappropriate in studies involving controversial/high-risk issues or vulnerable populations, such as minors	Respect for People
5	Using passive consent can lead to future problems for the study or the evaluator	NA
9	Other	
4	No explanation given	

(Table continues)

Respondents Who Judged the Evaluator's Actions as Definitely or Probably Not Ethically Problematic (n = 76)

Percentage	Explanation	Relevant Guiding Principle
43	Passive consent is an ethically acceptable procedure for obtaining informed consent	Respect for People
25	Passive consent is acceptable as long as it does not not focus on controversial/sensitive issues or expose participants to significant harm	Respect for People
16	Passive consent may be necessary in order to obtain a valid, representative sample	Systematic Inquiry
4	This scenario does not raise an ethical issue	NA
5	Other	NA
7	No explanation given	NA

Note. NA = Not Applicable

TABLE 8: Respondents' Explanations: The Revised Report Scenario

Respondents Who Judged the Evaluator's Actions as Definitely or Probably Ethically Problematic (n = 198)

Percentage	Explanation	Relevant Guiding Principle
61	Substantively altering a fair and balanced report undermines the accuracy, integrity, and scientific rigor of the evaluation	Integrity/Honesty
22	Deleting quotes is not an appropriate solution, however, it might be acceptable to modify the report in other ways	Integrity/Honesty
4	Altering the report violates the evaluator's primary responsibility, which is to the foundation (the client)	Welfare*, Integrity/Honesty
9	Other	NA
4	No explanation given	NA

(Table continues)

Respondents Who Judged the Evaluator's Actions as Definitely or Probably ***Not*** *Ethically Problematic (n = 126)*

Percentage	Explanation	Relevant Guiding Principle
70	As long as the report's key findings are not substantively altered with the quotes removed, the evaluator is behaving ethically	Integrity/Honesty
13	Evaluators have an ethical responsibility to be sensitive to the needs of programs and stakeholders, as well as to the political consequences of their reports	Respect for People, Welfare*
5	This scenario does not raise an ethical issue	NA
5	Other	NA
7	No explanation given	NA

Note. NA = Not Applicable

*Full title: "Responsibilities for General and Public Welfare"

TABLE 9: Respondents' Explanations: The Advisory Group Scenario

Respondents Who Judged the Evaluator's Actions as Definitely or Probably Ethically Problematic (n = 155)

Percentage	Explanation	Relevant Guiding Principle
30	Stakeholder participation generates input that is needed for an *accurate* evaluation	Systematic Inquiry
22	Stakeholder participation is a "given" in an ethical evaluation	Welfare*
21	It is unethical to form an "advisory" group and then not use them as such	Integrity/Honesty
17	The *usefulness utilization* of an evaluation is decreased if stakeholders aren't meaningfully involved	Integrity/Honesty
7	Other	NA
2	No explanation given	NA

(Table continues)

Respondents Who Judged the Evaluator's Actions as Definitely or Probably Not Ethically Problematic (n = 126)

Percentage	Explanation	Relevant Guiding Principle
50	This scenario raises a methodological or philosophical issue, not an ethical one	NA
15	The initial understanding between the evaluator and the stakeholders may not have provided for *extensive* stakeholder involvement	Integrity/Honesty
12	The advisory group does have the opportunity to provide *some* input into the evaluation	Integrity/Honesty
11	The evaluator is the expert; involving stakeholders "in depth" is not necessary, and might even compromise the objectivity of the evaluation	Systematic Inquiry
5	Other	NA
6	No explanation given	NA

Note. NA = Not Applicable

Full title: "Responsibilities for General and Public Welfare"

TABLE 10: Relationship of Combined Scenarios to Key Variables

	Number of Problematic Scenarios			
Variables	*Zero*	*1-2*	*3*	χ^2 *(df, n) or F(df)*
Roles (Mean)				
Consultant	5.74_a	5.30_a	4.77_b	4.56 (2, 381)**
Scientist	3.97	3.83	3.76	ns
Reporter	3.10	3.46	3.40	ns
Facilitator	4.16	4.49	4.26	ns
Guiding Principles				
Useful (Mean)	2.94_a	3.43_b	3.83_c	6.68 (2, 194)**
Familiar (%)	56	52	47	ns
Pol. views (Mean)	5.35	4.92	5.03	ns
Experience (Mean)				
Years (Non-PBC)	18.3_a	12.6_b	9.0_c	7.74 (2, 311)***
# evals. (Non-PBC)	4.1_a	3.6_a	2.9_b	8.24 (2, 312)***
% external (PBC)	90.0_a	81.0_a	52.2_b	4.82 (2, 70)**
% team	65.3	59.7	62.5	ns

(Table continues)

	Number of Problematic Scenarios			
Variables	*Zero*	*1-2*	*3*	χ^2*(df, n) or F(df)*
Degree (%)				
Doctoral	77	65	52	6.24 (2, 372)*
MA/BA	23	35	48	
Employment (%)				
PBC	37	18	14	8.07, (2, 391)**
Non-PBC	63	82	86	
Sex (%)				
Male	52	49	40	ns
Female	48	51	60	

Note. Means with different subscripts differ significantly at $p < .05$ or lower in the Tukey honestly significant difference comparison. PBC = Private Business/Consulting. Non-PBC = those employed in other settings. Means for "# evals." refer to survey scale values, not actual number of evaluations conducted.

*$p < .05$. **$p < .01$. ***$p < .001$.

1. Between 22% and 24% of the sample (depending on the scenario) chose not to offer predictions, sometimes writing that they "didn't have a clue" as to what the correct percentages might be.